Supplementary Information for

3d virtual histology of human cerebellum by x-ray phase-contrast tomography

Mareike Töpperwien, Franziska van der Meer, Christine Stadelmann and Tim Salditt

Tim Salditt.
E-mail: tsaldit@gwdg.de

This PDF file includes:

> Figs. S1 to S2
> Tables S1 to S2
> References for SI reference citations

SI Methods

Automatic cell segmentation. The automatic cell segmentation is based on the spherical Hough transform which finds spherical objects in a 3d array by building a voting matrix. Here, a slightly modified version of the Matlab implementation by Luke Xie [1, 2] was used. In order to decrease computation time, the data sets were divided into several sub-volumes. In the first step of building the voting matrix, a pre-voting of relevant pixels was performed based on the magnitude of the gradient in the image. In order to have the same conditions for all sub-volumes, the gradient threshold was set to an absolute value chosen based on visual inspection. As the gradient is largest towards the edge of the spheres this leads to a volume mainly consisting of the sphere contours. In the course of the algorithm, this will then later lead to maxima at the sphere center positions. In addition, the following parameters have to be specified: range of possible radii, the minimal intensity value of the detected sphere center, the multiple radii tolerance ($0.1 \rightarrow$ several possible radii, $1 \rightarrow$ only 'principal' radius), and the expected deviation from a perfect sphere. The latter parameter specifies the radius of a filter used in the search of local maxima in the accumulation array. The larger the deviation from a perfect sphere the larger the scattering of possible center positions in the accumulation matrix. In order for the algorithm to detect these multiple maxima in the accumulation array as a single sphere the filter radius has to be chosen larger so that all possible centers are combined to one (minimum filter radius is 3). As for the gradient threshold, the minimal center intensity value was chosen as an absolute value to create equal conditions throughout the entire volume. Based on these parameters, which are listed in Suppl. Table S2 for the two datasets, the algorithm finds all possible spheres in the 3d volumes. Note that the performance of the algorithm was found to improve by increasing the signal to noise ratio at the cost of resolution. For the synchrotron dataset, resampling by a factor of 2 as well as a Gaussian filter with a standard deviation of 0.3 pixels seemed to be a convenient choice, whereas in the case of the laboratory data the volume had to be reconstructed at full resolution with a subsequent lowpass filtering step, as in the resampled dataset, the cell radii were below 2 pixels and therefore not detectable. It also proved to be sufficient to select all relevant parameters based on a small sub-volume of the entire dataset which makes this process conveniently fast.

In order to avoid multiple segmentations of the same cell, mainly due to an overlap between the sub-volumes, automatically detected cells with a distance smaller than $2 \cdot (\min(\text{radius}) - 0.5)$ pixels were merged to one cell. The subtraction of 0.5 pixels from the minimum radius accounts for the fact that only integer values for the radii are determined by the algorithm.

Determination of the angular distribution of neighboring cells in the granular layer. In order to determine the angular distribution of nearest neighbors within the granular layer, the vectors between each cell and neighboring cells was determined. To this end, all cells within the first coordination shell were considered, as defined by an upper limit $r_{\max}$ given by the first minimum of the pair correlation function. Hence each cell has on average N nearest neighbors, where N is the coordination number. The position vectors were normalized to a length of 1 and their orientation was expressed in spherical coordi-

nates with the polar angle $\theta \in [0, 180]^\circ$ and azimuthal angle $\phi \in [0, 360]^\circ$ (cf. Fig. S1(a)). Note that prior to determining the orientation of the position vectors, the dataset was aligned by rotation such that the flat dendritic tree of the Purkinje cells lies approximately in the xy-plane, corresponding to a polar angle $\theta = 90^\circ$, and the y-axis is approximately parallel to the interface between the molecular and granular layer (cf. Fig. S1(b)). To determine cell densities as a function of the angles θ and ϕ, the spherical surface to which each vector points was divided into segments with constant surface area, leading to larger sampling steps in the polar angle θ towards the poles (cf. Fig. S2). The data was sampled with $N_\theta = 60$ points for the polar and $N_\phi = 120$ points for the azimuthal angle, respectively. By counting the number of cells within each of the spherical segments and dividing it by the corresponding surface area, the angular distribution of cell densities was computed. In order to visualize the angular distribution of nearest neighbors, the determined cell densities were plotted against the angles θ and ϕ both in a Cartesian as well as polar grid, as can be seen in Figs. 4 (main manuscript) and S2. As every cell within the granular layer is considered for the determination of position vectors to its nearest neighbors, each of the vectors occurs twice, though with changing sign. Hence, the upper and lower hemisphere are symmetric and it is sufficient to depict only one of the two hemispheres in the polar plot, in this case $0^\circ \leq \theta \leq 90^\circ$.
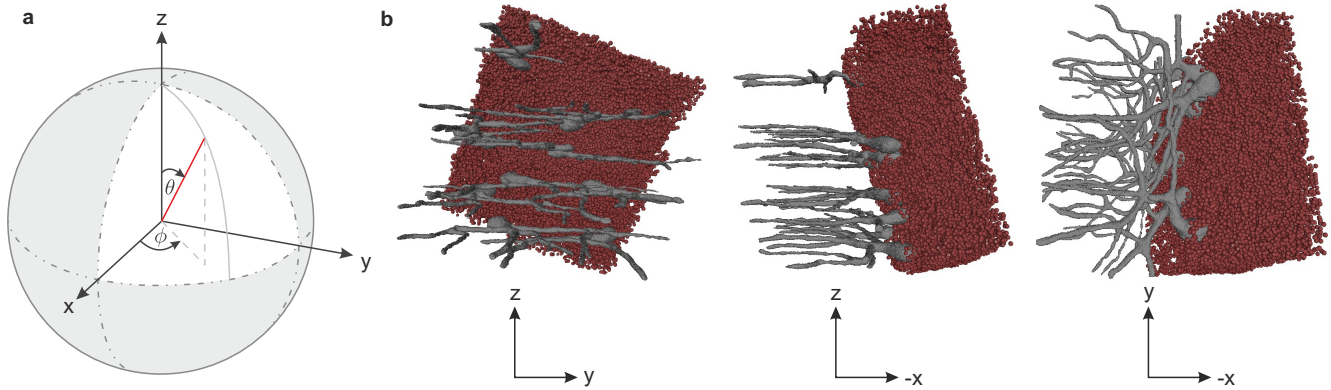
Mareike Töpperwien, Franziska van der Meer, Christine Stadelmann and Tim Salditt

Fig. S1. Visualization of the coordinate system for the determination of the anistropy in the cellular distribution within the first coordination shell. (a) The coordinate system is defined in spherical coordinates. Hence, the orientation of each of the position vectors between two adjacent cells, which was normalized to a length of 1, is given by means of the polar angle $\theta \in [0, 180]^{\circ}$ and the azimuthal angle $\phi \in [0, 360]^{\circ}$. (b) Prior to determining the orientation of the position vector between two cells, the dataset was aligned with respect to the Purkinje cell layer such that the flat dendritic tree lies approximately in the xy-plane, corresponding to an polar angle $\theta = 90^{\circ}$, and the y-axis is approximately parallel to the interface between the molecular and granular layer.
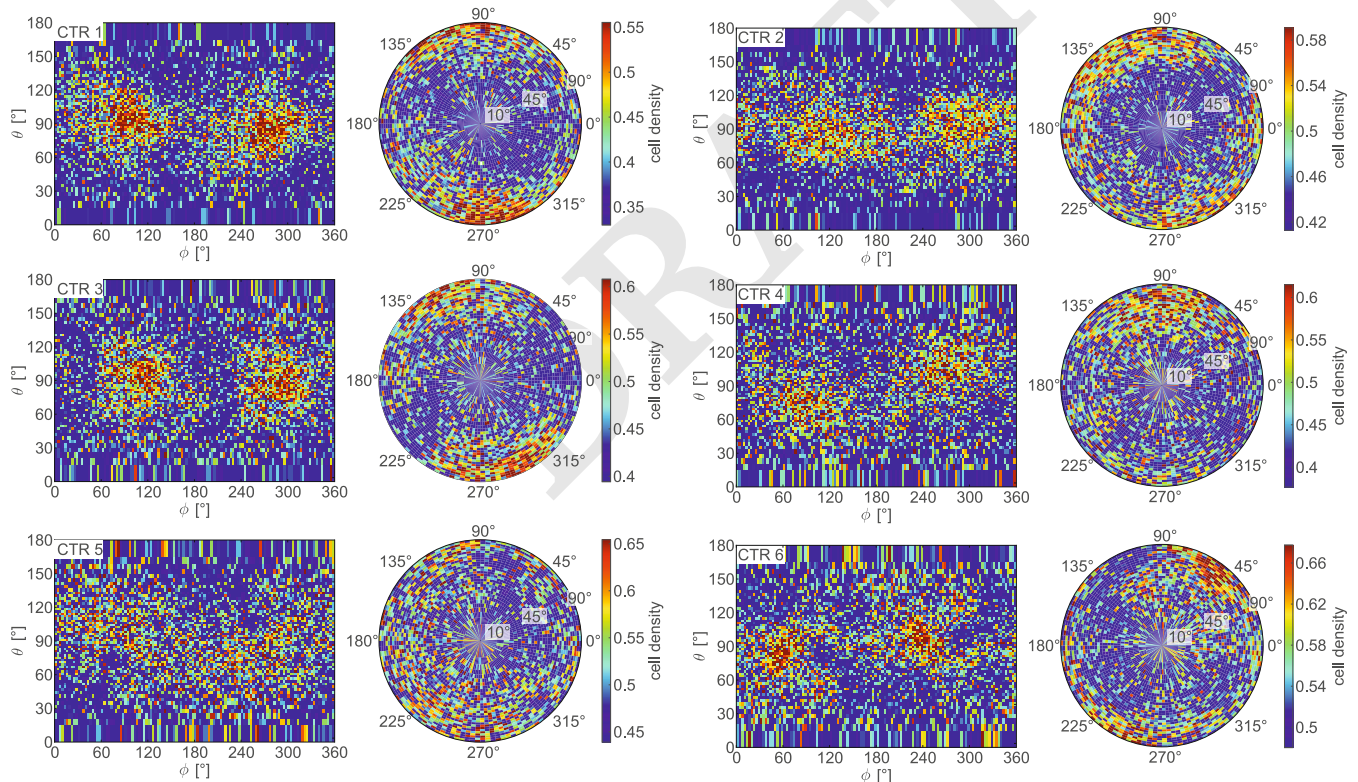

Fig. S2. Angular distribution of nearest neighbors in the first coordination shell determined for biopsy punches from six control patients, denoted as CTR 1 - CTR 6, which were imaged at the GINIX setup. Note that CTR 1 corresponds to the dataset shown in the main manuscript. All datasets were approximately aligned according to the geometry depicted in Fig. S15. In all patients, an anisotropic distribution of nearest neighbors can be observed as hotspots in the angular distribution of position vectors between each cell in the granular layer and its corresponding neighbors lying within the first coordination shell. Although this anisotropy is differently pronounced throughout the patients, each dataset shows that a major part of the cells is approximately arranged in parallel to the large dendritic tree of the Purkinje cells as these hotspots can be found at $\theta \simeq 90^{\circ}$, corresponding to the xy-plane.
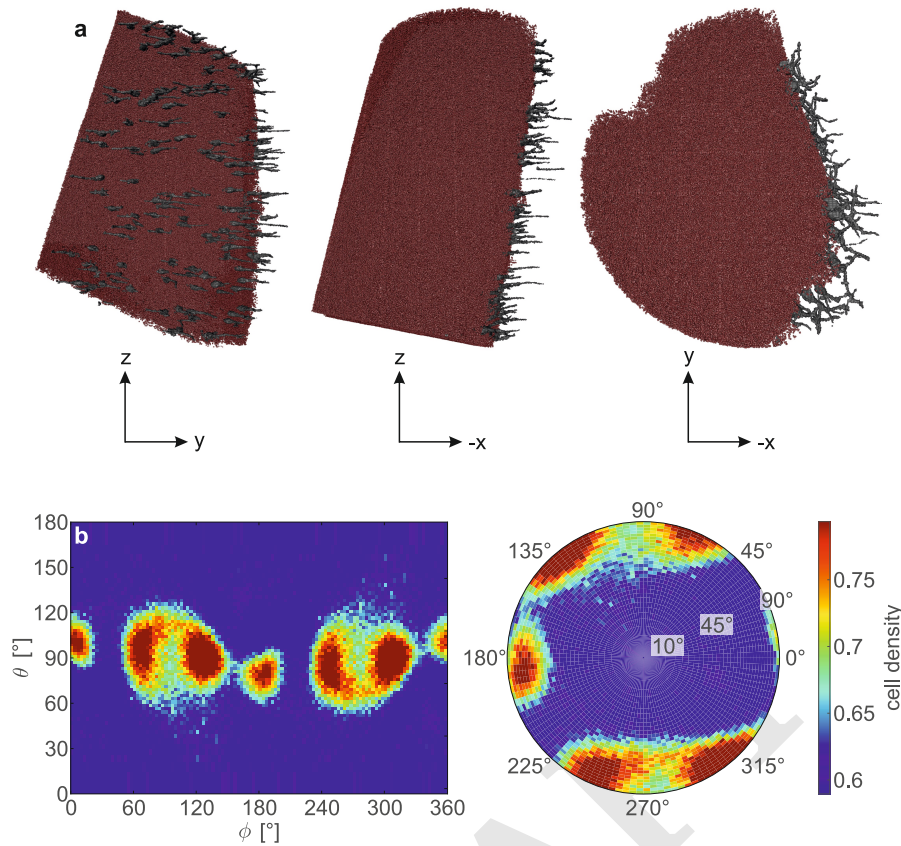
Fig. S3. Angular distribution of nearest neighbors in the granular layer of the laboratory dataset. (a) Analogous to the synchrotron dataset in Fig. S1, the dataset is aligned with respect to the Purkinje cell layer prior to the determination of the angular distribution of nearest neighbors. To this end, it was transformed such that the dendritic tree lies approximately in the xy-plane and hence at $\theta \simeq 90°$ (cf. Fig. S1 for the definition of the angles θ and ϕ). In contrast to the synchrotron dataset, the curvature of the interface between the granular and molecular layer prevents an alignment along its course and hence, the x- and y-axes were not specifically aligned with respect to the Purkinje cells. (b) Angular distribution of nearest neighbors in the granular layer of the laboratory dataset. The majority of nearest neighbors in the granular layer are clearly distributed in parallel to the dendritic tree of the Purkinje cells as hotspots in the angular distribution are visible at $\theta \simeq 90°$. In contrast to the synchrotron dataset (Fig. S2(a)), several of these hotspots distributed along the azimuthal angle ϕ can be recognized, indicating a larger number of main directions in the distribution of cells. This might be explained by the curvature of the granular layer, which is visible due to the larger field of view in the laboratory dataset.
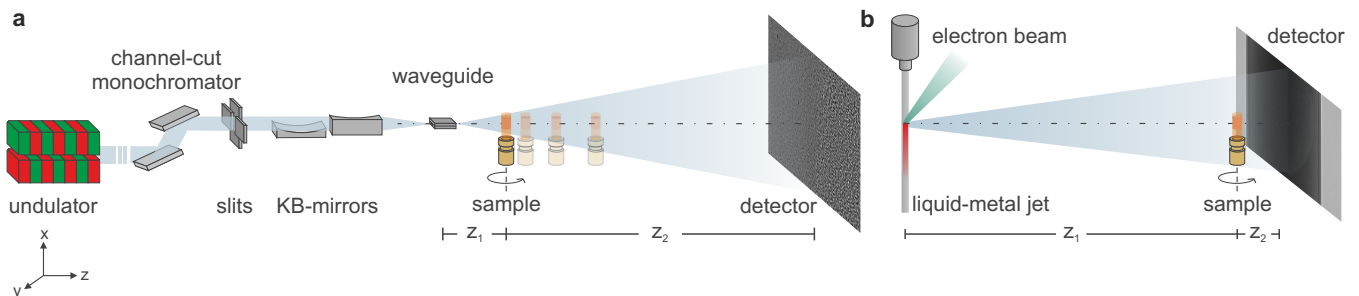

Fig. S4. Experimental setups. (a) Sketch of the GINIX endstation at beamline P10 of the PETRAIII storage ring at DESY. The monochromatized radiation is prefocused by a set of Kirkpatrick-Baez mirrors and is further reduced in diameter and increased in coherence by propagation in a x-ray waveguide placed in the focal plane. Subsequently the radiation reaches the sample which is mounted on a fully motorized stage. A holographic image is recorded by a scintillator-based sCMOS camera further downstream. (b) Sketch of the laboratory setup. The x-rays are generated by a liquid-metal jet microfocus source and are recorded by a high-resolution scintillator-based CCD camera. The sample is placed close to the detector in order to minimize source blurring effects.
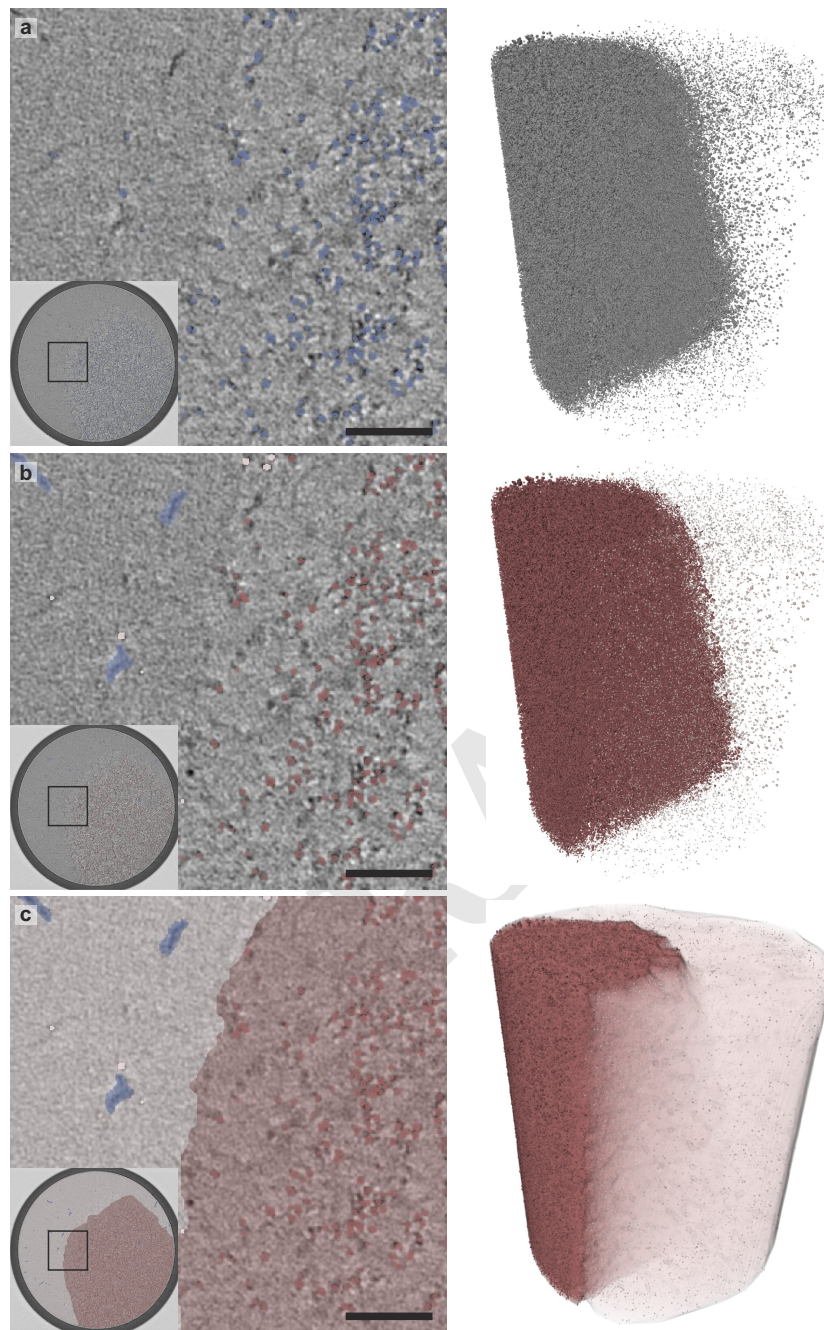
Mareike Töpperwien, Franziska van der Meer, Christine Stadelmann and Tim Saldittc

Fig. S5. Results of the automated segmentation procedure for the laboratory data. (a) Output of the segmentation algorithm without further processing, visualized by overlay in a transverse slice through the reconstructed 3d density, along with the corresponding volume rendering showing the 3d spatial distribution. For better visibility a $250 \times 250 \ \mu m^2$ inset is shown. (b) Automatic segmentation of the granular and molecular layer in the laboratory dataset, after manual removal of the vessels. As in the synchrotron dataset, the two layers are distinguished based on computing the mean distance to the 35 nearest neighbors of each cell. (c) Volume determined for the two layers within the cerebellum after post-processing of the segmentation results. Scale bars: 50 μm
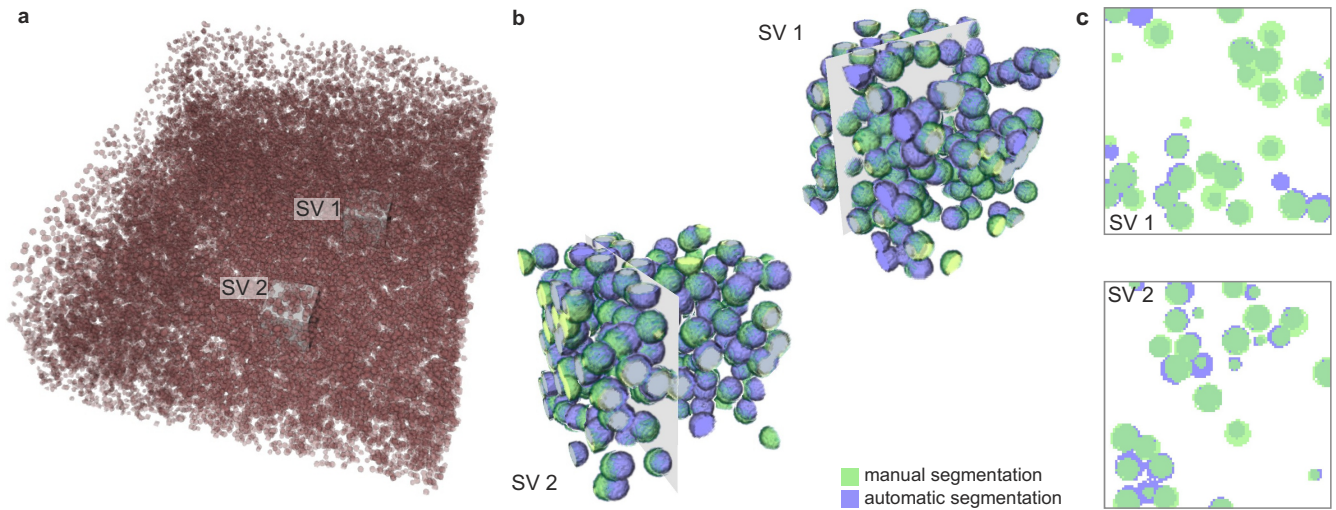
Fig. S6. Comparison between manual and automatic segmentation results for the the granular layer in the synchrotron dataset. (a) Volume rendering of the whole granular layer with the two randomly selected sub-volumes (SV 1 and SV 2 with volume 100^3 pixels) for which a manual segmentation was performed as a benchmark for the performance of the segmentation algorithm. (b) Comparison between the results of the automatic script and the manually obtained results. Manually detected cells are shown in green, automatically detected cells are depicted in blue. To evaluate the performance of the automatic segmentation procedure the precision and recall for each subvolume were determined, resulting in $(p, r) = (1, 1)$ in sub-volume 1 and $(p, r) = (0.994, 1)$ in sub-volume 2. This indicates that that algorithm performs almost perfectly in the region of the granular layer.
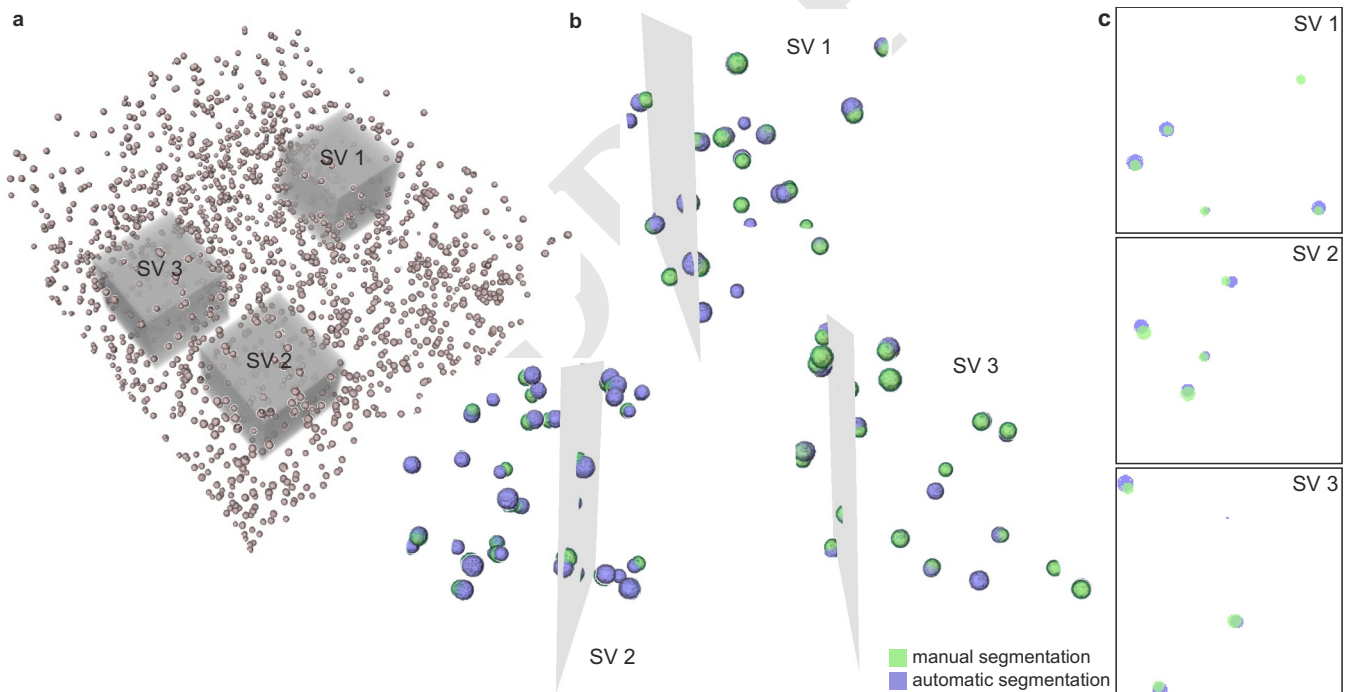

Fig. S7. Comparison between manual and automatic segmentation results for the the molecular layer in the synchrotron dataset. As for the granular layer (shown in Fig. S6) the performance of the algorithm is judged with respect to a manual segmentation which is considered to be the ground truth. (a) Volume rendering of the molecular layer with the three randomly chosen sub-volumes used for the manual segmentation. Due to the lower cell density and therefore reduced statistics, three sub-volumes with twice the side length compared to granular layer (200^3 pixels) were chosen. (b) Comparison between the manually segmented (green) and automatically detected cells (blue). Compared to the results of the granular layer, the number of actual false positives is higher which leads to slightly worse precision and recall values within the sub-volumes ($(p, r) = (0.9, 1)$ in sub-volume 1, $(p, r) = (0.88, 1)$ in sub-volume 2 and $(p, r) = (0.9, 0.95)$ in sub-volume 3). The slightly worse performance of the algorithm within the molecular layer can be attributed to the diversity in nuclei size and shape compared to the more uniform distribution within the granular layer (see Fig. S10).
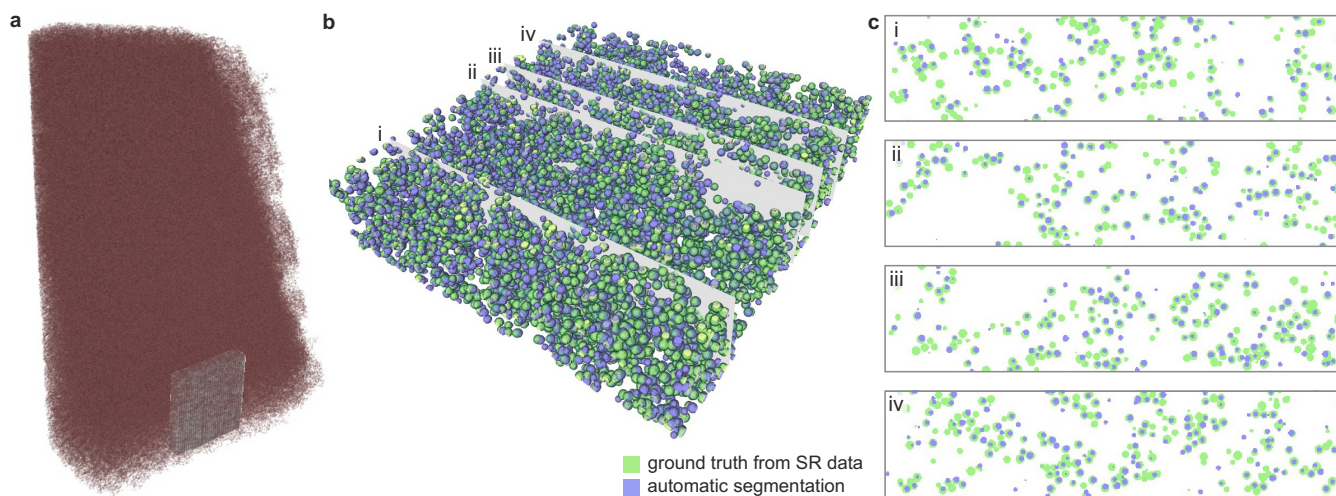
Mareike Töpperwien, Franziska van der Meer, Christine Stadelmann and Tim Salditt

Fig. S8. Evaluation of the performance of the automatic segmentation algorithm for the granular layer in the laboratory data, with the results from the synchrotron data taken as the ground truth. (a) Volume rendering of the granular layer with a box indicating the region of the synchrotron data taken for comparison. (b) Comparison between the synchrotron results (green) with the segmentation results obtained for the laboratory data (blue). (c) Slices through the rendering of the results from the synchrotron and the laboratory datasets for better visibility. The positions of the slices are indicated in b. Precision and recall are $(p, r) = (0.85, 0.93)$ showing that despite the lower data quality obtained by the laboratory setup, the data quality is still sufficient to automatically detect the vast majority of cells in the granular layer of the cerebellum.
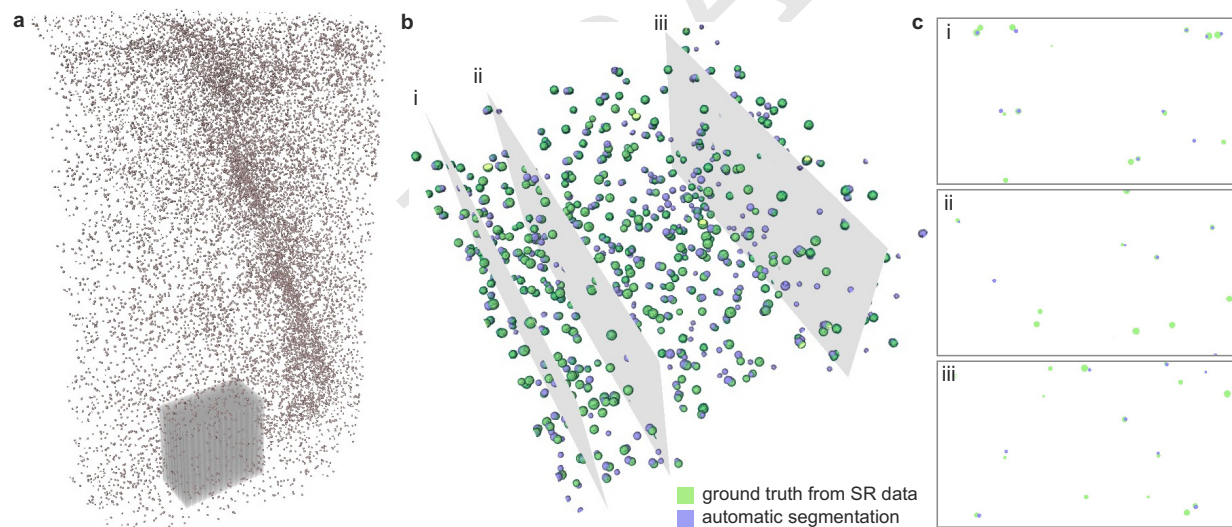

Fig. S9. Evaluation of the performance of the automatic segmentation algorithm for the molecular layer in the laboratory data. Note that manual segmentation in the laboratory data turned out to be very challenging, and in fact much less convincing than the automatically obtained results from the synchrotron data. Therefore, the synchrotron result is again taken as ground truth. (a) Volume rendering of the molecular layer with a box indicating the region of the synchrotron data taken for comparison. (b) Comparison between the segmentation results obtained from the synchrotron data (green) and from the laboratory data (blue). (c) Again slices through these segmentation results are shown for better visibility. Precision and recall are $(p, r) = (0.71, 0.72)$, indicating that the chosen parameters in the segmentation algorithm work well for the granular layer but are not perfectly suited for the molecular layer. Possibly, two different sets of parameters, each optimized with respect to the segmentation of one of the two layers, could improve the results.
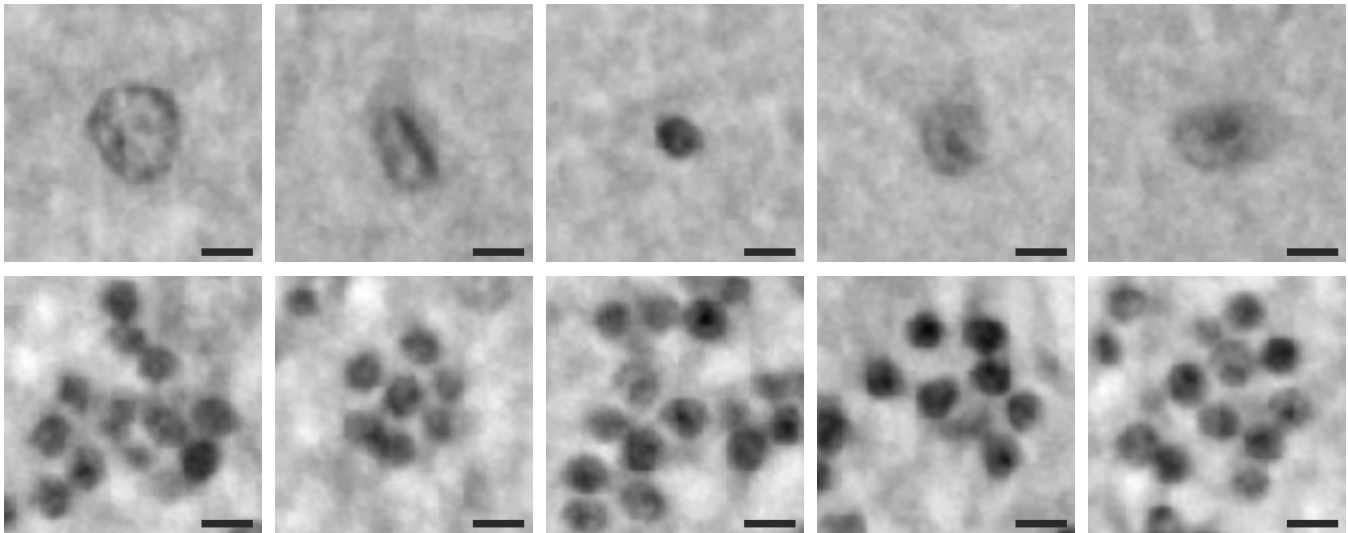
Fig. S10. Comparison of cell types found in the molecular and granular layer of the cerebellum (synchrotron dataset). Upper row: Examples of five cells in the molecular layer, showing the diversity of cell shapes and sizes within this layer. Lower row: Examples of cells in the granular layer. Compared to the cells from the molecular layer, cell shapes and sizes appear to be uniform throughout the layer and the the overall cell size is considerably smaller than in the molecular layer, which makes automatic segmentation of all cells with one set of parameters challenging. Scale bars: 5 µm
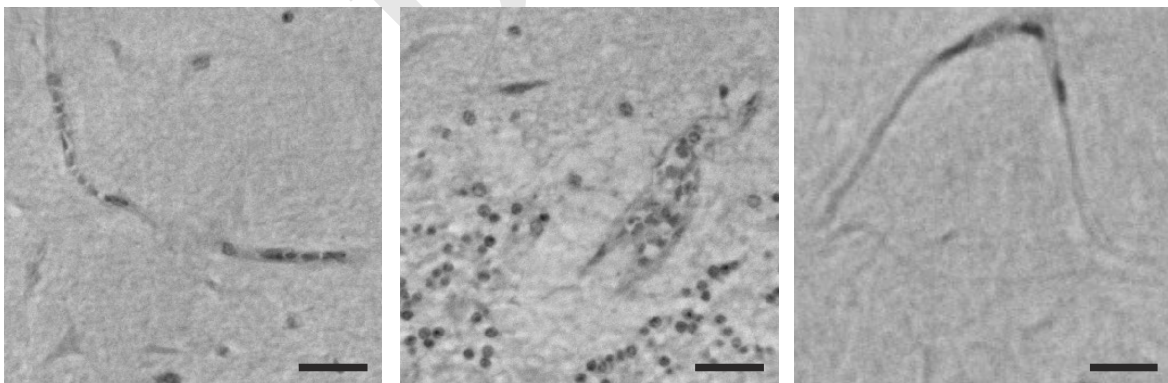

Fig. S11. Blood vessels found in the cerebellum (synchrotron dataset). The blood cells within the vessels are well resolved and therefore also found by the automatic cell segmentation algorithm. Scale bars: 25 µm (left and middle) and 15 µm (right)

Mareike Töpperwien, Franziska van der Meer, Christine Stadelmann and Tim Salditt
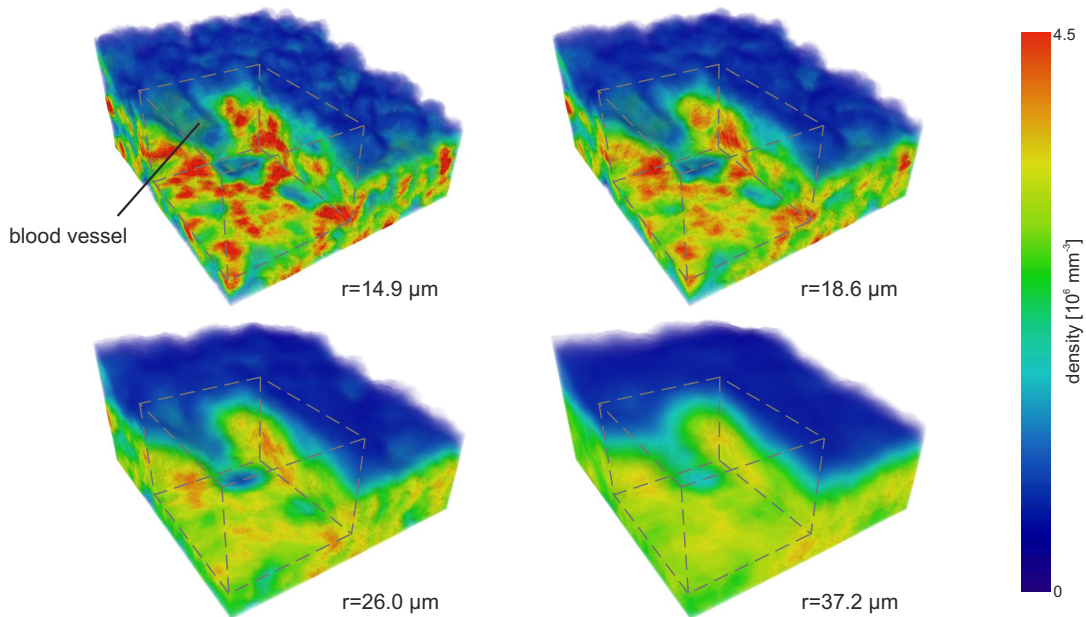
Fig. S12. Local cell density distribution within the granular layer of the cerebellum as obtained from the synchrotron dataset. The local density is computed by a convolution between the cell center positions and a sphere with varying radius. The example with a sphere radius of 14.9 µm (top left) clearly shows the clustering of cells within this layer as hotspots of cell density can be clearly recognized. With increasing sphere radius (top right to bottom right) the differences in cell density are vanishing, leading to an almost uniform density distribution within the granular layer. However, towards the interface between the granular and molecular layer (at the top of the shown volume) cell density still becomes considerably smaller, indicating that the transition between those two layers is not sharp but that the cell density gradually decreases towards the low-cellular molecular layer. Note that for better visibility of the inner structures, transparency is higher for less dense regions in all images.
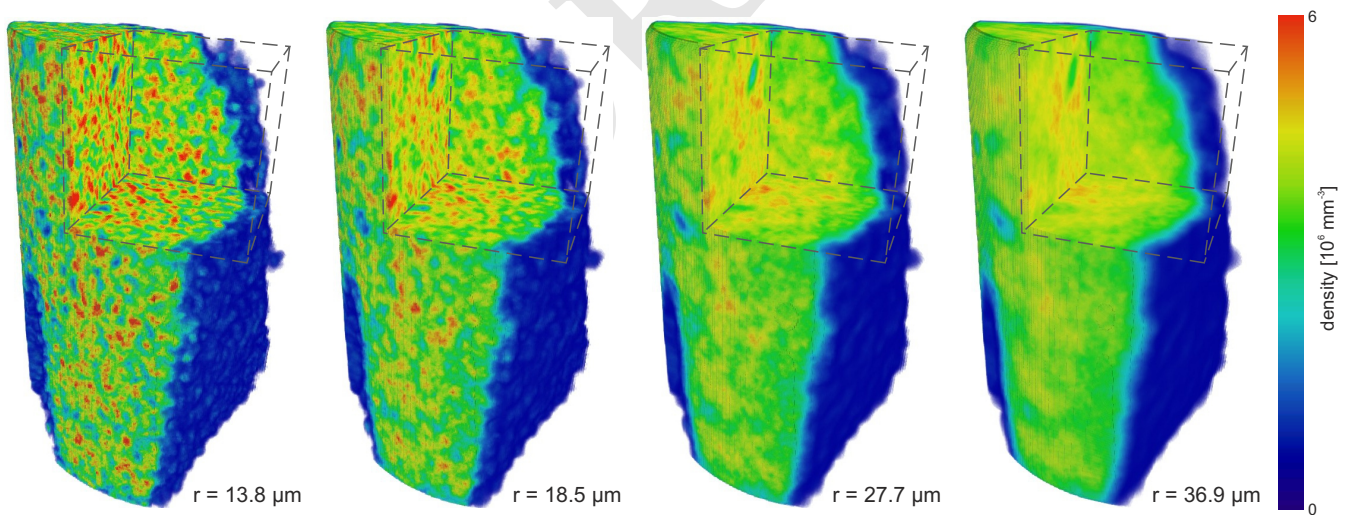

Fig. S13. Local cell density distribution within the granular layer of the cerebellum as obtained from the laboratory dataset. The same behavior as in the synchrotron dataset can be recognized, namely the clustering of cells within the granular layer which can be clearly recognized in the left panel and which vanishes with increasing sphere radius. It is also evident that the density of the granular layer becomes larger towards the center of the layer. This not only indicates the smooth transition between the molecular and granular layer, as in the synchrotron dataset, but also that the synchrotron dataset, which was obtained at the interface between these two layer, contains a less dense part of the granular layer, explaining the difference in overall density between the two datasets.
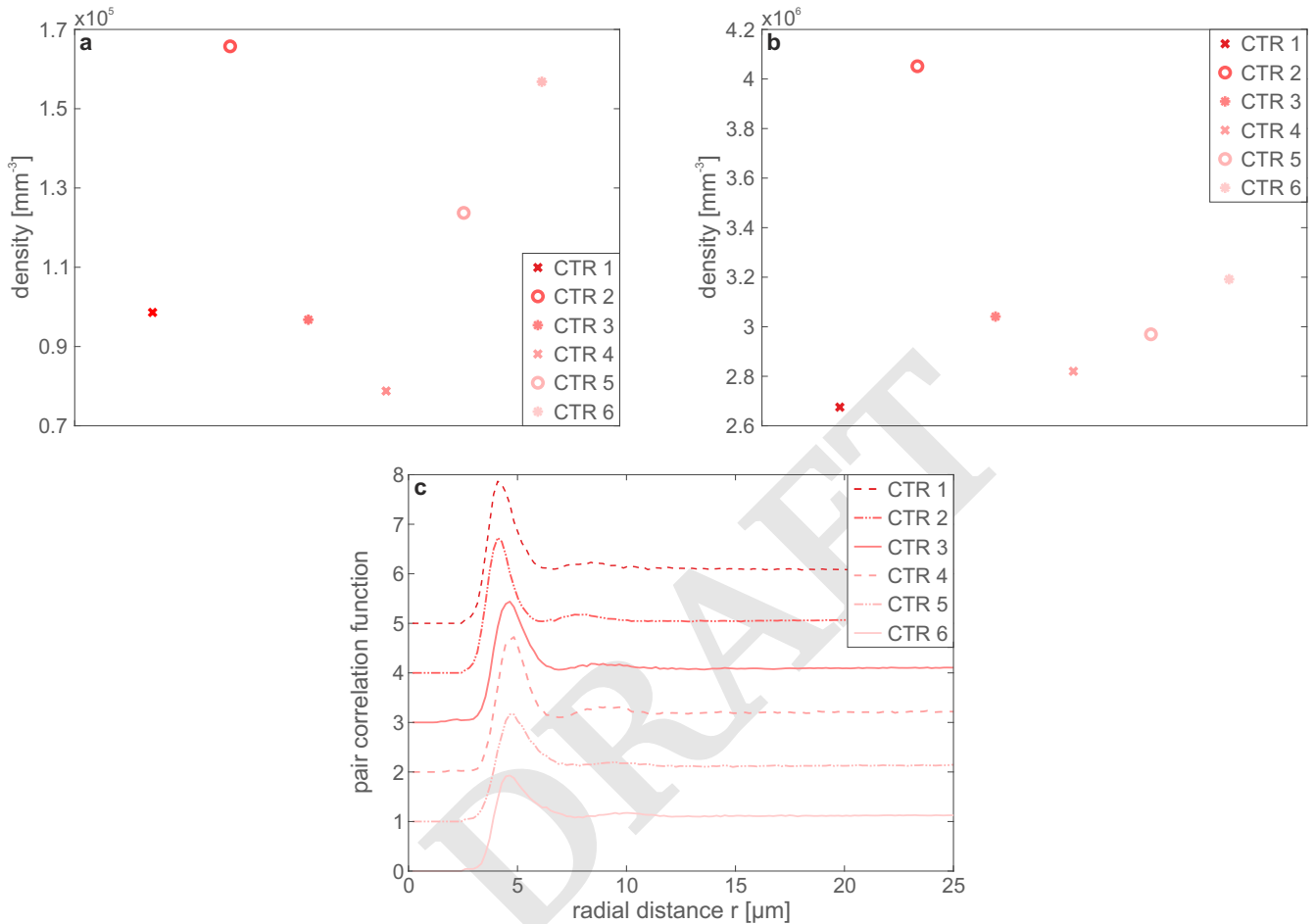
Fig. S14. Comparison between the determined cell densities and pair correlation functions for paraffin-embedded human cerebellum obtained from six control patients measured at the GINIX setup. Note that patient CTR 1 corresponds to the data shown in the main text. For all other samples, data acquisition and analysis as well as post-processing steps were carried out analogously to the procedures described in the Materials and Methods section. (a) Cell densities in the molecular layer. For better visibility each data point was shifted laterally compared to its predecessor. The mean density value yields $\rho_{ML} = (1.2 \pm 0.2) \cdot 10^5$ mm^{-3} (standard error of the mean (SEM)). (b) Cell densities in the granular layer. Compared to the molecular layer, the densities are considerably higher, yielding a mean value of $\rho_{GL} = (3.1 \pm 0.2) \cdot 10^6$ mm^{-3} (SEM). (c) Pair correlation functions for the granular layer. For better visibility, all curves were vertically shifted by 1 compared to their predecessor. The general shape of the pair correlation function is similar in all patients, showing two distinct maxima at mean peak positions of 4.4 ± 0.1 μm and 9.1 ± 0.3 μm (SEM), corresponding to approximately once and twice the mean cell diameter (4.3 ± 0.1 μm (SEM)). The mean number of nearest neighbors of each cell, determined via the first coordination number, yields 5.9 ± 0.3 (SEM).

Mareike Töpperwien, Franziska van der Meer, Christine Stadelmann and Tim Salditt
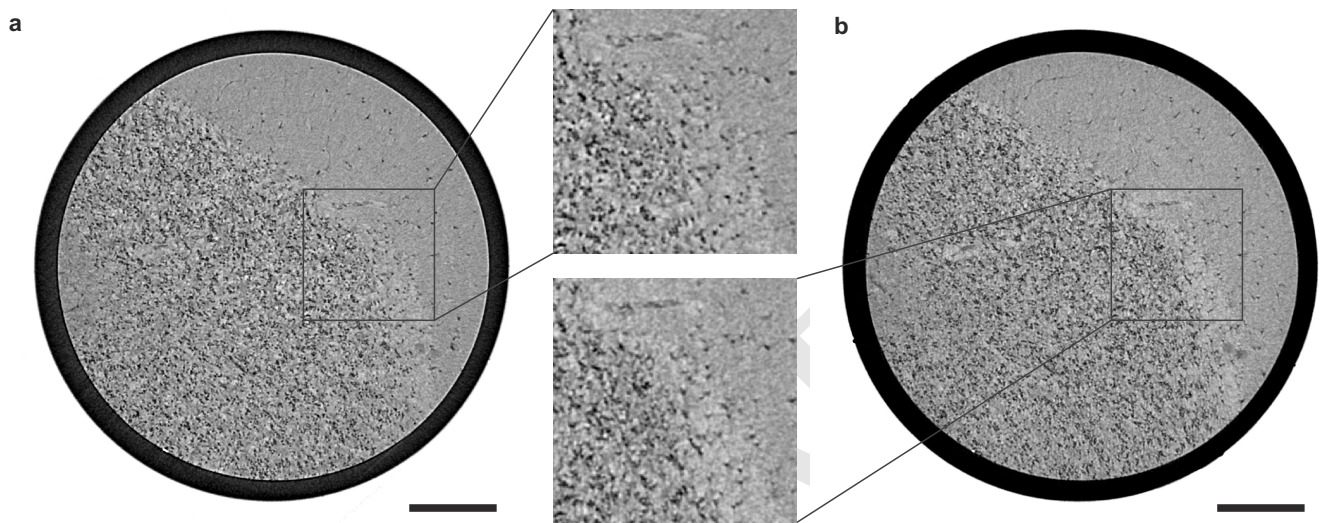
Fig. S15. Comparison between the results obtained by a liquid-metal jet and a rotating anode microfocus source. (a) Virtual slice through the reconstructed volume obtained by a liquid-metal jet source and the parameters specified in the main text. The rectangle marks the position of the magnified region shown in the center on the top. Within the depicted slice, the tissue is resolved at cellular detail, enabling the automatic segmentation of over a million of cells within the reconstructed volume. (b) Corresponding virtual slice through the reconstructed volume obtained by a rotating anode source equipped with a copper target. Apart from the source, the setup is analogous to the setup of the liquid-metal jet source with a scintillator-based lens-coupled CCD detector with a pixel size of 550 nm. The minimum source size lies in the range of 70 μm (30 mA, 40 kV), still enabling sufficient partial coherence for phase-contrast effects to occur but requiring a larger source-to-detector distance and hence smaller magnification to minimize source blurring effects. The source-to-sample distance was $z_1 = 456$ mm while the sample-to-detector distance was set to $z_2 = 9$ mm, leading to a magnification of 1.02 and an effective pixel size of 530 nm. Experimental parameters were the same as those in the main text. The data analysis was performed analogously to the liquid-metal jet setup with regularization parameters $\alpha = 0.06$ and $\gamma = 0.21$ used for phase retrieval. As in the case of the liquid-metal jet source, single cells can be resolved throughout the reconstructed volume despite the lower coherence due to the larger source spot. The region marked by the rectangle is depicted at higher magnification in the center at the bottom, showing that compared to the liquid-metal jet source, the reconstructed density seems slightly blurred, but single cells can still be resolved. This makes us confident that the automatic segmentation of over a million of cells can also be carried out with a rotating anode microfocus source.

Table S1. Experimental parameters

	Synchrotron CT	Laboratory CT
Detector	sCMOS	CCD
Pixel size p [µm]	6.5	0.54
z_1 [mm]	145	158.75
z_2 [mm]	4901	27.05
eff. pixel size p_{eff} [nm]	186	461
field of view [mm^2]	0.38×0.38	1.53×1.15 (h×v)
Energy (spectrum)/K_α [keV]	8	0-40/9.25
e-beam power [W]		57
exposure time [s]	0.05	50
number of projections	1500	1000
number of distances	4	1
angular range [°]	0-180	0-180
BAC, α		0.01
BAC, γ		0.55
CTF, δ/β	50	
CTF, $\lim_2$	0	
half-period resolution (FWHM)*	370 nm	1.3 µm

* The resolution was estimated via the edge steepness of an edge between selected cells from the granular layer and the surrounding tissue. The given value is half of the FWHM of an error function fitted to this edge.

Table S2. Table containing the parameters chosen for the spherical Hough transform in order to automatically detect all cells in the reconstructed volumes

	radius range [pixels]	gradient threshold	filter radius [pixels]	multiple radius tolerance	object center intensity threshold
Synchrotron dataset	3 − 10	$3.4 \cdot 10^{-4}$	3	1	0
Laboratory dataset	1 − 6	$3.1 \cdot 10^{-5}$	3	1	$1.6 \cdot 10^{-4}$

The radius range sets the minimum and maximum value of radii for the detected spheres. Spheres with radii lying outside of this range are not detected. Therefore, without knowledge of the approximate nuclear sizes, this range should be relatively large. However, this comes at the cost of longer computation times. The gradient threshold sets the threshold for pixels being considered for calculating the accumulation matrix. The filter radius accounts for non-perfect spheres. The larger the deviation, the larger this value has to be chosen, with 3 being the minimal filter radius. The multiple radius tolerance determines the number of possible radii for each sphere center position. It ranges from 0.1 (highest tolerance and therefore maximum number of radii) to 1 (lowest tolerance, only the 'principal' radius is selected). The object center intensity threshold gives the minimum gray value which the detected center pixels should have in order to be considered as the center of a sphere.

Mareike Töpperwien, Franziska van der Meer, Christine Stadelmann and Tim Salditt
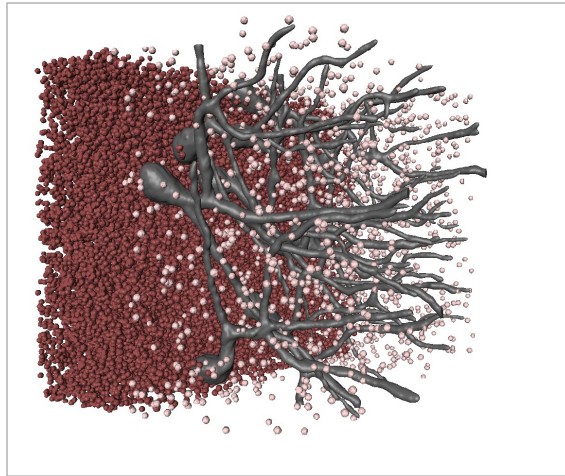
Movie S1. Animation showing virtual slices through the reconstructed 3d volume obtained at the synchrotron in the approximate orientation of the Purkinje cell layer, revealing the dendritic tree of each Purkinje cell as well as the cellular distribution in the granular and molecular layer at high detail. The performance of the automatic cell segmentation is visualized by an overlay of the obtained results on the corresponding virtual slices through the volume. Note that in the edge regions of the reconstructed volume no cells are automatically segmented, as they were excluded in the analysis. Finally, the results of the automatic segmentation in the ML and GL as well as the semi-automatic segmentation of the Purkinje cell layer are depicted in 3d, enabling the visualization of cellular distributions within the cerebellum.
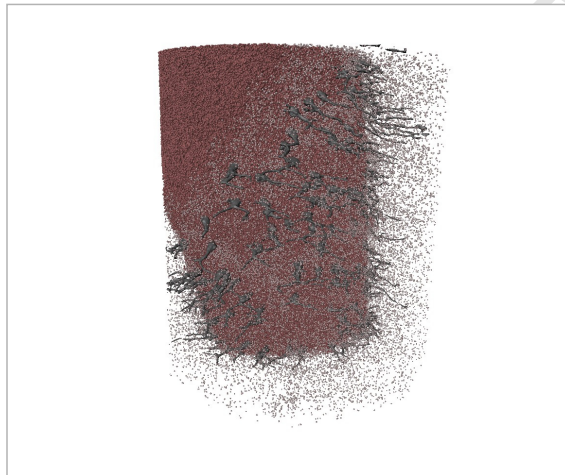
Movie S2. Animation of the laboratory dataset, also depicting virtual slices through the reconstructed 3d density. Again, the Purkinje cells and cells in the molecular and granular layer are resolved in high detail. For the visualization of the segmentation, the automatically determined cell positions are overlayed on virtual slices through the 3d volume. A first impression of the 3d density distribution is given by a volume rendering, in which background voxels are rendered transparent and dense structures in dark red, whereas less dense regions are depicted in light red. Finally, the automatically segmented cells in the ML and GL as well as the semi-automatically segmented Purkinje cells are again visualized in 3d. For two selected cells in the Purkinje cell layer a manually refined segmentation is shown.

References

1. Xie L (2014) Spherical hough transform for 3d images (https://de.mathworks.com/matlabcentral/fileexchange/48219-spherical-hough-transform-for-3d-images).
2. Peng T, Balijepalli A, Gupta SK, LeBrun T (2007) Algorithms for on-line monitoring of micro spheres in an optical tweezers-based assembly cell. *Journal of computing and information science in engineering* 7(4):330–338.